

Working Paper 94-51
Statistics and Econometrics Series 19
December 1994

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

NON-PARAMETRIC SPECIFICATION TESTING OF NON-NESTED ECONOMETRIC MODELS

Miguel A. Delgado, Qi Li and Thanasis Stengos*

Abstract

We consider the non-nested testing problem of non-parametric regressions. We show that, when the regression functions are unknown under both the null and the alternative hypotheses, an extension of the J-test procedure of Davidson and Mackinnon (1981) will lead to a test statistic with well defined asymptotic properties. The derivation of the test statistic involves double kernel estimation. Monte Carlo simulations suggest that the test has good size and power characteristics.

Key Words

Non-parametric Test; Non-nested Models; Double Kernel Estimation.

*Delgado, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid; Li, Department of Economics, University of Guelph; Stengos, Department of Economics, University of Guelph. The first author would like to thank the "Dirección General de Investigación Científica y Técnica" of Spain, reference number: PB92-0247, for research support. The second and the third authors are grateful to the financial support from the Social Sciences and Humanities Research Council of Canada.

Non-parametric Specification Testing of Non-nested Econometric Models.*

Miguel A. Delgado
Departamento de Econometria y Estadistica
Universidad Carlos III
Calle Madrid 126
28903 Getafe, Madrid, Spain

Qi Li
Department of Economics
University of Guelph
Guelph, Ontario N1G 2W1
Canada

Thanasis Stengos
Department of Economics
University of Guelph
Guelph, Ontario N1G 2W1
Canada

April 1995

Abstract

We consider the non-nested testing problem of non-parametric regressions. We show that, when the regression functions are unknown under both the null and the alternative hypotheses, an extension of the J-test procedure of Davidson and MacKinnon (1981) will lead to a test statistic with well defined asymptotic properties. The derivation of the test statistic involves double kernel estimation. Monte Carlo simulations suggest that the test has good size and power characteristics.

Key Words: Non-parametric test; Non-nested models; Double kernel estimation.

*The first author would like to thank the "Direccion General de Investigacion Cientifica y Tecnica" of Spain, reference number: PB92-0247, for research support. The second and the third authors are grateful to the financial support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

This paper proposes a testing procedure for discriminating between alternative sets of regressors in a non-parametric context. The literature on non-parametric testing of regression models is quite extensive. Non-parametric methods have been used for specification testing of a parametric model against a non-parametric alternative, see Eubank and Spiegelman (1990), Hall and Hart (1990), Hong and White (1991), Wooldridge (1992), Härdle and Mammen (1993), Whang and Andrews (1993), Horowitz and Härdle, de Jong and Bierens (1994), Fan and Li (1994), and Delgado and Stengos (1994), to mention only a few.

Discriminating between non-nested sets of regressors is a well motivated problem. Existing tests assume a particular functional form of the regression function and are consistent in the direction of precisely parameterized alternatives, see Cox (1961, 1962), Pesaran (1974), Davidson and MacKinnon (1981) and Fisher and McAleer (1981), also see MacKinnon (1992) for a survey. Recently Delgado and Stengos (1994) have proposed an extension of the J-test of Davidson and MacKinnon that is consistent against non-parametric alternatives. The above test still assumes a particular parametric regression curve under the null hypothesis. Hence, it is still not robust to functional mis-specification. In this paper, we propose to test a non-parametric regression model in the direction of non-parametric non-nested alternative. As in Delgado and Stengos (1994), the present paper is an extension of the J-test to the non-parametric environment. The proposed test relies on the use of double kernel estimation and is, after multiplied by the \sqrt{n} factor, normally distributed.

In the next section we present a consistent non-parametric test for non-nested models; the proofs of the main results are collected in Appendix A. Section 3 presents results of some Monte Carlo simulations.

2 The Test Statistic and Its Asymptotic Distribution

Data consists of independent observations $\{(x_i, z_i, y_i), i = 1, \dots, n\}$ identically distributed as the $R^p \times R^q \times R$ -valued multivariate random variable (x, z, y) . The researcher faces the

problem of choosing between the alternative sets of explanatory variables x and z which are non-nested, in the sense that none of the σ -algebras corresponding to x and z are contained in the other. Our selection rule is based in a non-parametric model specification procedure. Under the null hypothesis $H_0: E[y|x, z] = m(x)$, we have

$$y_i = m(x_i) + u_i, \quad i = 1, \dots, n, \quad (1)$$

where $E(u_i|x_i, z_i) = 0$. Under the alternative hypothesis $H_a: E[y|x, z] = m_a(z)$, we have

$$y_i = m_a(z_i) + \epsilon_i, \quad i = 1, \dots, n, \quad (2)$$

where $E(\epsilon_i|x_i, z_i) = 0$. Proceeding in the way as suggested by Davidson and MacKinnon (1981), we nest the null and the alternative hypothesis in the artificial nested regression model,

$$y_i = (1 - \alpha)m(x_i) + \alpha m_a(z_i) + \eta_i,$$

where α is the nesting parameter and $\eta_i = (1 - \alpha)u_i + \alpha\epsilon_i$ is the composite error. The null hypothesis is reformulated as $H_0: \alpha = 0$, and the alternative as $H_a: \alpha = 1$.

The analog of the J-regression of Davidson and MacKinnon (1981) in the case of the two non-parametric regression functions examined here is given by

$$y_i = \theta(x_i) + \alpha \hat{m}_a(z_i) + \eta_i^*, \quad (3)$$

where $m_a(z)$ is replaced by its non-parametric kernel estimator $\hat{m}_a(z_i) = \hat{E}(y_i|x_i)$, $\theta(x_i) = (1 - \alpha)m(x_i)$ and $\eta_i^* = \eta_i + \alpha(m_a(z_i) - \hat{m}_a(z_i))$ is the composite error.

Following Robinson's (1988) semi-parametric estimation approach, α in (3) could be estimated by

$$\hat{\alpha} = \frac{\sum_i (\hat{m}_a(z_i) - \hat{E}(\hat{m}_a(z_i)|x_i))(y_i - \hat{E}(y_i|x_i))}{\sum_i (\hat{m}_a(z_i) - \hat{E}(\hat{m}_a(z_i)|x_i))^2},$$

where $\hat{E}(\cdot|x)$ is a nonparametric estimate of $E(\cdot|x)$. Our proposed test statistic is based on the numerator of $\hat{\alpha}$, given by

$$T_n = \frac{1}{n} \sum_i (\hat{m}_a(z_i) - \hat{E}(\hat{m}_a(z_i)|x_i))(y_i - \hat{E}(y_i|x_i)).$$

A direct application of Robinson's (1988) method in (3) requires two trimming parameters (in addition to the two smoothing parameters) to overcome the random denominator problem that arises in kernel estimation. Moreover, the technical difficulties of using a trimming method in the context of double kernel estimation prove difficult to overcome. Therefore, we choose to estimate a density-weighted relationship to avoid the random denominator problem, see Powell et al (1989). A density-weighted approach leads to a wider range of choices of smoothing parameters than the trimming method and the regularity conditions needed are weaker. We estimate $E(y_i|x_i)f(x_i)$ by

$$\hat{y}_i = \frac{1}{(n-1)a^p} \sum_{j \neq i} y_j K_{ij}, \quad (4)$$

and $f(x_i)$, the probability density function (p.d.f) of x_i , by $\hat{f}(x_i) = \frac{1}{(n-1)a^p} \sum_{j \neq i} K_{ij}$, where $K_{ij} = K(\frac{x_i - x_j}{a})$ is the kernel function and a is the smoothing parameter. We use a product kernel, $K(u) = \prod_{l=1}^p k(u_l)$; u_l is the l th component of u .

We estimate $m_a(z_i)f_a(z_i) = E(y_i|z_i)f_a(z_i)$ by

$$\tilde{y}_i = \frac{1}{(n-1)b^q} \sum_j y_j \bar{K}_{ij}, \quad (5)$$

and $f_a(z_i)$ is estimated by $\tilde{f}_a(z_i) = \frac{1}{(n-1)b^q} \sum_{j \neq i} \bar{K}_{ij}$, where $\bar{K}_{ij} = \bar{K}(\frac{z_i - z_j}{b})$ is the kernel function associated with z and b is the corresponding smoothing parameter ($\bar{K}(\cdot)$ is a product kernel). We also need to estimate $E\{m_a(z_i)f_a(z_i)|x_i\}f(x_i) = E\{[E(y_i|z_i)f_a(z_i)]|x\}f(x)$. Its kernel estimate is given by

$$\hat{\tilde{y}}_i = \frac{1}{(n-1)a^p} \sum_{j \neq i} \tilde{y}_j K_{ij} \equiv \frac{1}{(n-1)a^p} \sum_{j \neq i} \left[\frac{1}{(n-1)b^q} \sum_{l \neq j} y_l \bar{K}_{lj} \right] K_{ij}. \quad (6)$$

Our density-weighted test statistic will be based on

$$I_n = \frac{1}{n} \sum_i (\tilde{y}_i \hat{f}_i - \hat{\tilde{y}}_i)(y_i \hat{f}_i - \hat{y}_i), \quad (7)$$

which is, roughly speaking, a sample analogue of $I = E\{[\gamma(z) - E(\gamma(z)|x)]f(x)[y - E(y|x)]f(x)\}$, where $\gamma(z) = E(y|z)f_a(z)$.

To derive the asymptotic distribution of I_n , the following definitions and assumptions will be used. Let \mathcal{G}_l^α denote the class of functions such that if $g \in \mathcal{G}_l^\alpha$ ($\alpha > 0$ and $l \geq 1$

is an integer), then g is l times differentiable, g and its derivatives (up to order l) are all bounded by some function that has α th order finite moments. Also \mathcal{K}_l , $l \geq 1$, denote the class of even functions $k : \mathbb{R} \rightarrow \mathbb{R}$ satisfying $\int k(u)u^m du = \delta_{0m}$ for $m = 0, 1, \dots, l-1$ and $k(u) = O((1 + |u|^{l+1+\delta})^{-1})$, some $\delta > 0$. Denote $g(z) = E(y|z)$, $\gamma(z) = g(z)f_a(z)$ and $\xi = E(\gamma(z)|x)$.

(A1) (y_i, x_i, z_i) are independently distributed as (y, x, z) , x admits a pdf $f(x) \in \mathcal{G}_\nu^\infty$, also $m(x) \in \mathcal{G}_\nu^4$ and $\xi(x) \in \mathcal{G}_\nu^4$, where $\nu \geq 2$ is a positive integer. $\sigma^2(x)f(x)$ is uniformly bounded, where $\sigma^2(x) = E(u^2|x)$. Moreover z admits a pdf $f_a(z) \in \mathcal{G}_\mu^\infty$, $g(z) \in \mathcal{G}_\mu^4$, $m_a(z) \in \mathcal{G}_1^4$ and $E(m_a(z)|x) \in \mathcal{G}_1^4$, where $\mu \geq 2$ is a positive integer.

(A2) $k \in \mathcal{K}_\nu$. As $n \rightarrow \infty$, $na^{2p} \rightarrow \infty$, $na^{4\nu} \rightarrow 0$; $\bar{k} \in \mathcal{K}_\mu$. As $n \rightarrow \infty$, $nb^{2q} \rightarrow \infty$ and $nb^{4\mu} \rightarrow 0$.

Assumption (A1) presents some smoothness and moments conditions. (A2) is similar to the conditions used by Robinson (1988) or Fan, Li and Stengos (1992). It requires a higher order kernel to be used for $k(\bar{k})$ if $p \geq 4$ ($q \geq 4$).

The following theorems justify I_n as an asymptotic test. The proofs are presented in the appendix.

Theorem 1 Under assumptions (A1) and (A2), if H_0 is true, then as $n \rightarrow \infty$,

$J_n \xrightarrow{d} N(0, 1)$, where $J_n = \sqrt{n}I_n/\sqrt{\hat{\sigma}_0^2}$, $\hat{\sigma}_0^2 = \frac{1}{n} \sum (\hat{y}_i \hat{f}_i - \hat{\hat{y}}_i)^2 (y_i \hat{f}_i - \hat{y}_i)^2$ is a consistent estimator for $\sigma_0^2 = E\{[g(z_1) - E(g(z_1)|x_1)]^2 f^2(x_1) u_1^2 f^2(x_1)\}$, where $g(z_1) = E(y_1|z_1)$.

Theorem 2 Under assumptions (A1) and (A2), if H_a is true, then as $n \rightarrow \infty$,

$\text{Prob}[|J_n| > c] \rightarrow 1$ for any positive constant c .

3 Monte Carlo Results

In this section we investigate the small sample performance of the test statistic J_n in the context of some Monte Carlo experiments. We take the null model and the alternative model as given by $H_0: y_i = \beta_0 + (\beta_1 x_{1i} + \beta_2 x_{2i})^2 + u_{0i}$ and $H_a: y_i = \gamma_0 + (\gamma_1 z_{1i} + \gamma_2 z_{2i})^2 + u_{1i}$ respectively. The parameters $\beta_0, \beta_1, \beta_2, \gamma_0, \gamma_1, \gamma_2$ are all set to unity. The x 's and u 's are generated as

independent normal variates $N(0, 1)$. We generate the z 's as $z_{li} = \rho x_{li} + v_{li}$ ($l = 1, 2$), where v_{li} is distributed independently as $N(0, 1)$. By varying ρ we control the correlation coefficients between z_i 's and x_i 's. Similar Monte Carlo designs have been considered by Davidson and MacKinnon (1982), Godfrey and Pesaran (1983) and Delgado and Stengos (1994).

The sample size was chosen as $n = 50, 100, 200$ and the number of replications is 2000. We used a second order normal kernel and we set the bandwidth as $a_l = cx_{l, sd}n^{-1/6}$ and $b_l = cz_{l, sd}n^{-1/6}$, where $x_{l, sd}$ ($z_{l, sd}$) is the sample standard deviation of x_l (z_l , $l = 1, 2$) and c is a constant. We chose $c = 0.8, 1.0, 1.2$. Table 1 reports the size results. It can be seen that the results for our non-parametric test statistic are quite encouraging. In samples as small as $n = 50$, the size performance of the proposed test is quite good. Also the estimated size is not sensitive to the different c values used. For comparison purposes we also report the conventional J and JA test statistics that are computed based on the assumptions that the null and the alternative models are linear in x 's and z 's respectively.

The power results are presented in Table 2. The powers for different c values are virtually the same, hence we will only report the results for $c = 1$ to save space. Again the proposed non-parametric test performs quite well. As expected the test is asymptotically powerful. Also different values of ρ does not seem to affect the results considerably. Note that as ρ increases the correlation between the x 's and z 's increases as well and hence the two models become less distinguishable. As expected we also observe that the conventional J and JA tests are inconsistent, the number of rejections for both the J and the JA tests decreases as n increases.

The limited Monte Carlo results suggest that the proposed test performs adequately with respect to its size and power characteristics especially when compared to the traditional J and JA tests.

Table 1

Size Results: Proportion of Rejections when H_0 is true.

	$\lambda = 0.1$		$\lambda = 0.4$		$\lambda = 0.7$	
	5%	1%	5%	1%	5%	1%
$n = 50$						
$NP(c = 0.8)$	0.054	0.006	0.052	0.005	0.056	0.009
$NP(c = 1.0)$	0.050	0.005	0.051	0.003	0.060	0.009
$NP(c = 1.2)$	0.043	0.006	0.046	0.005	0.065	0.008
J	0.180	0.065	0.159	0.055	0.125	0.033
JA	0.065	0.019	0.068	0.019	0.061	0.020
$n = 100$						
$NP(c = 0.8)$	0.048	0.010	0.061	0.009	0.053	0.009
$NP(c = 1.0)$	0.046	0.005	0.055	0.007	0.055	0.009
$NP(c = 1.2)$	0.040	0.004	0.042	0.007	0.060	0.013
J	0.164	0.052	0.133	0.031	0.114	0.035
JA	0.060	0.011	0.062	0.010	0.060	0.015
$n = 200$						
$NP(c = 0.8)$	0.056	0.014	0.057	0.009	0.058	0.016
$NP(c = 1.0)$	0.051	0.012	0.054	0.009	0.053	0.013
$NP(c = 1.2)$	0.049	0.010	0.047	0.007	0.060	0.016
J	0.146	0.045	0.133	0.029	0.110	0.025
JA	0.053	0.009	0.056	0.014	0.057	0.011

Table 2

Power Results: Proportion of Rejections when H_a is true.

	$\lambda = 0.1$		$\lambda = 0.4$		$\lambda = 0.7$	
	5%	1%	5%	1%	5%	1%
$n = 50$						
NP	0.260	0.062	0.275	0.068	0.24	0.060
J	0.373	0.203	0.348	0.197	0.301	0.133
JA	0.173	0.081	0.142	0.073	0.140	0.060
$n = 100$						
NP	0.511	0.219	0.477	0.223	0.401	0.165
J	0.364	0.182	0.326	0.176	0.301	0.131
JA	0.143	0.061	0.144	0.070	0.132	0.053
$n = 200$						
NP	0.840	0.634	0.754	0.523	0.651	0.421
J	0.341	0.162	0.310	0.160	0.290	0.130
JA	0.139	0.060	0.137	0.070	0.130	0.050

Appendix A

Similar to (4) -(6), for any sequence $\{A_i, i = 1, \dots, n\}$, we define $\hat{A}_i = \frac{1}{(n-1)ap} \sum_{j \neq i} A_j K_{ij}$, $\tilde{A}_i = \frac{1}{(n-1)bp} \sum_{j \neq i} A_j \bar{K}_{ij}$, and $\hat{\tilde{A}}_i = \frac{1}{(n-1)ap} \sum_{j \neq i} [\frac{1}{(n-1)bp} \sum_{l \neq j} A_l \bar{K}_{jl}] K_{ij}$; they are kernel estimators for $E(A_i|x_i)f(x_i)$, $E(A_i|z_i)f_a(z_i)$ and $E[E(A_i|z_i)f_a(z_i)|x_i]f(x_i)$ respectively. Define $w(x_i, z_i) = m(x_i) - g(z_i)$. We will write $m_i = m(x_i)$, $g_i = g(z_i)$, $\gamma_i = \gamma(z_i)$, $\xi_i = \xi(x_i)$, $w_i = m_i - g_i$, $f_i = f(x_i)$ and $f_{ai} = f_a(z_i)$. There should be no confusion about these because when we define a function, it is clear from the context whether its argument is x , z or (x, z) .

Proof of Theorem 1

We first give a decomposition of I_n under H_0 . Recall that $g_i = E(y_i|z_i)$ and $w_i = m_i - g_i$. We have, under H_0 , $y_i = m_i + u_i = g_i + w_i + u_i$. Thus

$$\begin{aligned} y_i &= m_i + u_i, & \hat{y}_i &= \hat{m}_i + \hat{u}_i, \\ \tilde{y}_i &= \tilde{m}_i + \tilde{u}_i = \tilde{g}_i + \tilde{w}_i + \tilde{u}_i = \gamma_i + (\tilde{g}_i - \gamma_i) + \tilde{w}_i + \tilde{u}_i \\ \hat{\tilde{y}}_i &= \hat{\tilde{m}}_i + \hat{\tilde{u}}_i = \hat{\tilde{g}}_i + \hat{\tilde{w}}_i + \hat{\tilde{u}}_i = \hat{\gamma}_i + (\hat{\tilde{g}}_i - \hat{\gamma}_i) + \hat{\tilde{w}}_i + \hat{\tilde{u}}_i \end{aligned} \quad (A.1)$$

where $\gamma_i = g_i f_{ai}$. Substitute (A.1) into (7), we get

$$I_n = I_{1n} + I_{2n},$$

where

$$I_{1n} = \frac{1}{n} \sum_i [\gamma_i \hat{f}_i - \hat{\gamma}_i] [(m_i \hat{f}_i - \hat{m}_i) + u_i \hat{f}_i - \hat{u}_i],$$

and

$$I_{2n} = \frac{1}{n} \sum_i [(\tilde{g}_i - \gamma_i) \hat{f}_i + \tilde{w}_i \hat{f}_i + \tilde{u}_i \hat{f}_i - (\hat{\tilde{g}}_i - \hat{\gamma}_i) - \hat{\tilde{w}}_i - \hat{\tilde{u}}_i] [(m_i \hat{f}_i - \hat{m}_i) + u_i \hat{f}_i - \hat{u}_i].$$

We will prove $\sqrt{n}I_n \xrightarrow{d} N(0, \sigma_0^2)$ by showing that $\sqrt{n}I_{1n} \xrightarrow{d} N(0, \sigma_0^2)$ and $I_{2n} = o_p(n^{-1/2})$.

Proof of $\sqrt{n}I_{1n} \xrightarrow{d} N(0, \sigma_0^2)$:

Recall that $\xi_i = E[\gamma_i|x_i]$ and define $v_i = \gamma_i - \xi_i$. Denote $S_{Af,Bf} = \frac{1}{n} \sum_i A_i \hat{f}_i B_i \hat{f}_i$, $S_{A,B} = \frac{1}{n} \sum_i A_i B_i$, $S_{Af,B} = \frac{1}{n} \sum_i A_i \hat{f}_i B_i$. Also $S_A = S_{A,A}$. Then I_{1n} can be written as

$$I_{1n} = \frac{1}{n} \sum_i [(\xi_i \hat{f}_i - \hat{\xi}_i) + v_i \hat{f}_i - \hat{v}_i] [(m_i \hat{f}_i - \hat{m}_i) + u_i \hat{f}_i - \hat{u}_i]$$

$$\begin{aligned} &\equiv S_{(\xi f - \hat{\xi}) + v f - \hat{v}, (m f - \hat{m}) + u f - \hat{u}} = S_{u f, v f} + \{S_{\xi f - \hat{\xi}, m f - \hat{m}} + S_{\xi f - \hat{\xi}, u f} - S_{\xi f - \hat{\xi}, \hat{u}} \\ &\quad + S_{v f, m f - \hat{m}} - S_{v f, \hat{u}} - S_{\hat{v}, m f - \hat{m}} - S_{\hat{v}, u f} + S_{\hat{v}, \hat{u}}\} \end{aligned} \quad (A.2)$$

All the terms inside the curly bracket of (A.2) are $o_p(n^{-1/2})$ by lemma B.4. $\sqrt{n}(S_{u f, v f} - S_{u f, v f}) = O_p(n^{-1}a^{-p} + a^{2\nu}) = o_p(1)$ by proposition 7 of Fan, Li and Stengos (1992). Hence $\sqrt{n}I_{1n} = \sqrt{n}S_{u f, v f} + o_p(1) \rightarrow N(0, \sigma_0^2)$ by the Levi-Linderberg central limit theorem.

Proof of $I_{2n} = o_p(n^{-1/2})$

$$\begin{aligned} I_{2n} &= \frac{1}{n} \sum_i [(\tilde{g}_i - \gamma_i)\hat{f}_i + \tilde{w}_i\hat{f}_i + \tilde{u}_i\hat{f}_i - (\hat{g}_i - \hat{\gamma}_i) - \hat{w}_i - \hat{u}_i][(\tilde{m}_i\hat{f}_i - \hat{m}_i) + u_i\hat{f}_i - \hat{u}_i] \\ &= S_{(\tilde{g}-\gamma)\hat{f} + \tilde{w}\hat{f} + \tilde{u}\hat{f} - (\hat{g}-\hat{\gamma}) - \hat{w} - \hat{u}, (m\hat{f} - \hat{m}) + u\hat{f} - \hat{u}} \end{aligned}$$

We need to show that $S_{(\tilde{g}-\gamma)\hat{f}, m\hat{f}-\hat{m}}, S_{(\tilde{g}-\gamma)\hat{f}, u\hat{f}}, S_{(\tilde{g}-\gamma)\hat{f}, \hat{u}}, S_{\tilde{g}-\hat{\gamma}, m\hat{f}-\hat{m}}, S_{\tilde{g}-\hat{\gamma}, u\hat{f}}, S_{\tilde{g}-\hat{\gamma}, \hat{u}}, S_{\tilde{r}\hat{f}, m\hat{f}-\hat{m}}, S_{\tilde{r}\hat{f}, u\hat{f}}, S_{\tilde{r}\hat{f}, \hat{u}}, S_{\tilde{r}, m\hat{f}-\hat{m}}, S_{\tilde{r}, u\hat{f}}, S_{\tilde{r}, \hat{u}}$ are all $o_p(n^{-1/2})$, where $r = w$ or u . It suffices to show that (i) $S_{(\tilde{g}-\hat{\gamma})\hat{f}} = o_p(n^{-1/2})$, (ii) $S_{\tilde{g}-\hat{\gamma}} = o_p(n^{-1/2})$, (iii) $S_{(\tilde{g}-\hat{\gamma})\hat{f}, u\hat{f}} = o_p(n^{-1/2})$, (iv) $S_{\tilde{g}-\hat{\gamma}, u\hat{f}} = o_p(n^{-1/2})$, (v) $S_{\tilde{r}\hat{f}} = o_p(n^{-1/2})$, (vi) $S_{\tilde{r}} = o_p(n^{-1/2})$, (vii) $S_{\tilde{r}\hat{f}, u\hat{f}} = o_p(n^{-1/2})$, (viii) $S_{\tilde{r}, u\hat{f}} = o_p(n^{-1/2})$, ($r = w$ or u). Because (i)-(viii), together with lemma B.4 and the Cauchy inequality imply that all the other terms in I_{2n} are also $o_p(n^{-1/2})$, we prove (i)-(viii) in propositions 1 to 8 below.

Proposition 1 $S_{(\tilde{g}-\gamma)\hat{f}} = O((nb^q)^{-1} + b^{2\mu})$.

Proof: First by adding and subtracting term, we have $(\tilde{g}_i - \gamma_i)\hat{f}_i = (\tilde{g}_i - g_i\tilde{f}_{ai})\hat{f}_i + g_i(\tilde{f}_{ai} - f_{ai})\hat{f}_i \triangleq F_i + H_i$, where $F_i = (\tilde{g}_i - g_i\tilde{f}_{ai})\hat{f}_i$ and $H_i = g_i(\tilde{f}_{ai} - f_{ai})\hat{f}_i$. Thus we have $S_{(\tilde{g}-\gamma)\hat{f}} \equiv S_{F+H}$. By Cauchy inequality, it suffices to show that $S_F = o_p(n^{-1/2})$ and $S_H = o_p(n^{-1/2})$.

We prove these below.

Proof of $S_F = O((nb^q)^{-1} + b^{2\mu})$:

$E|S_F| = \frac{1}{n} \sum E(F_i^2) = E(F_1^2)$. Using $F_1 = (\tilde{g}_1 - g_1\tilde{f}_a(z_1))\hat{f}_1 = [((n-1)b^q)^{-1} \sum_{i \neq 1} (g_i - g_1)\tilde{K}_{1i}][((n-1)a^p)^{-1} \sum_{j \neq 1} K_{1j}]$, we have

$$E|S_F| = E\{[((n-1)b^q)^{-1} \sum_{i \neq 1} (g_i - g_1)\tilde{K}_{1i}][((n-1)b^q)^{-1} \sum_{i' \neq 1} (g_{i'} - g_1)\tilde{K}_{1i'}]\}$$

$$[((n-1)a^p)^{-1} \sum_{j \neq 1} K_{1j}] [((n-1)a^p)^{-1} \sum_{j' \neq 1} K_{1j'}] \quad (\text{A.3})$$

If i, i', j, j' are all different from each other, then conditional on 1, the terms in different square brackets in (A.3) are independent, using lemmas B.1 and B.3. In this case, (A.3) is

$$\begin{aligned} & E\{[b^{-q}E_1(g_2 - g_1)\bar{K}_{12}][b^{-q}E_1(g_3 - g_1)\bar{K}_{13}][a^{-p}E_1(K_{14})][a^{-p}E_1(K_{15})]\} \\ & \leq E\{[D(z_1)b^\mu][D(z_1)b^\mu][O(1)][O(1)]\} = O(b^{2\mu}). \end{aligned}$$

Next we consider the case that i, i', j, j' (in (A.3)) take three different values. There are six different combinations, but symmetry reduces it to three different cases. (1) $i = i'$, (2) $j = j'$ and (3) $i = j$, or $i = j'$, or $i' = j$, or $i' = j'$. We first consider case (1): $i = i'$ and in this case, using lemmas B.1 and B.3, (A.3) becomes

$$\begin{aligned} & E\{[((n-1)b^q)^{-2} \sum_{i \neq 1} (g_i - g_1)^2 \bar{K}_{1i}^2] [((n-1)a^p)^{-1} \sum_{j \neq 1} K_{1j}] [((n-1)a^p)^{-1} \sum_{j' \neq 1} K_{1j'}]\} \\ & = E\{[n^{-1}b^{-2q}E_1(g_2 - g_1)^2 \bar{K}_{12}^2][a^{-p}E_1(K_{13})][a^{-p}E_1(K_{14})]\} \\ & = O((nb^q)^{-1})O(1)O(1) = O((nb^q)^{-1}) \end{aligned}$$

For case (2): $j = j'$, and (A.3) becomes (using lemmas B.1 and B.3)

$$\begin{aligned} & E\{[((n-1)b^q)^{-1} \sum_{i \neq 1} (g_i - g_1) \bar{K}_{1i}] [((n-1)b^q)^{-1} \sum_{i' \neq 1} (g_{i'} - g_1) \bar{K}_{1i'}] [((n-1)a^p)^{-2} \sum_{j \neq 1} K_{1j}^2]\} \\ & = E\{[b^{-q}E_1(g_2 - g_1)\bar{K}_{12}][b^{-q}E_1(g_3 - g_1)\bar{K}_{13}][na^{-2p}E_1(K_{14}^2)]\} \\ & \leq E\{[D(z_1)b^\mu][D(z_1)b^\mu][n^{-1}a^{-q}O(1)]\} = O((na^p)^{-1}b^{2\mu}) = o(b^{2\mu}) \end{aligned}$$

For case (3), the two subcases are symmetric, we only consider $j' = i$, then (A.3) becomes

$$\begin{aligned} & E\{[((n-1)a^p)^{-1}((n-1)b^q)^{-1} \sum_{i \neq 1} (g_i - g_1) \bar{K}_{1i} K_{1i}] [((n-1)b^q)^{-1} \sum_{j \neq 1} (g_j - g_1) \bar{K}_{1j}] [((n-1)a^p)^{-1} \sum_{i' \neq 1} K_{1i'}]\} \\ & = E\{[((n-1)a^p)^{-1}b^{-q}E_1(g_2 - g_1)\bar{K}_{12}K_{12}][b^{-q}E_1(g_3 - g_1)\bar{K}_{13}][a^{-p}E_1(K_{14})]\} \\ & \leq E\{[n^{-1}D(z_1)b^\mu O(1)][D(z_1)b^\mu][O(1)]\} = O(n^{-1}b^{2\mu}) = o(b^{2\mu}). \end{aligned}$$

where we used $E_1[(g_2 - g_1)\bar{K}_{12}K_{12}] = E_1[(g(z_2) - g(z_1))\bar{K}_{12}E_{1,z_2}(K(\frac{z_1 - z_2}{a}))] = E_1[(g_2 - g_1)\bar{K}_{12}O(a^p)] \leq D(z_1)b^{q+\mu}a^p$ by lemma B.3.

When i, i', j, j' take no more than two values, we will have at most two summations, and using similar arguments as above, it is easy to show that, in this case, (A.3) will have a

smaller order. Hence $E|S_F| = O((nb^q)^{-1} + b^{2\mu})$.

Proof of $S_H = O((nb^q)^{-1} + b^{2\mu})$:

$E|S_H| = \frac{1}{n} \sum_i E(H_i^2) = E(H_1^2)$. Using $H_1 = g(z_1)(\tilde{f}_a(z_1) - f_a(z_1))\hat{f}_1$, we have

$$\begin{aligned} E|S_H| &= E\left\{[g(z_1)n^{-1} \sum_{i \neq 1} (b^{-q} \bar{K}_{1i} - f_a(z_1))][g(z_1)n^{-1} \sum_{i' \neq 1} (b^{-q} \bar{K}_{1i'} - f_a(z_1))]\right. \\ &\quad \left. [((n-1)a^p)^{-1} \sum_{j \neq 1} K_{1j}] [((n-1)a^p)^{-1} \sum_{j' \neq 1} K_{1j'}]\right\} \end{aligned} \quad (A.4)$$

Comparing (A.4) with (A.3), the main difference is that $(\tilde{g}_1 - g_1 \tilde{f}_a(z_1))$ in (A.3) is replaced by $(\tilde{f}_a(z_1) - f_a(z_1))$ here. This amounts to replacing lemma B.3 by lemma B.2 in the proof. Hence (A.4) has the same order as (A.3), and $S_H = O_p((nb^q)^{-1} + b^{2\mu})$.

Proposition 2 $S_{\hat{g}-\hat{\gamma}} = O((nb^q)^{-1} + b^{2\mu})$. **Proof:** By adding and subtracting terms, we can rewrite \hat{g}_i as

$$\begin{aligned} \hat{g}_i &\equiv \frac{1}{(n-1)a^p} \sum_{j \neq i} \left[\frac{1}{(n-1)b^q} \sum_{l \neq j} g(z_l) \bar{K}_{jl} \right] K_{ij} \\ &= \frac{1}{(n-1)a^p} \sum_{j \neq i} g(z_j) f_a(z_j) K_{ij} + \frac{1}{(n-1)a^p} \sum_{j \neq i} g(z_j) ((\tilde{f}_a(z_j) - f_a(z_j)) K_{ij} \\ &\quad + \frac{1}{(n-1)a^p} \sum_{j \neq i} \left[\frac{1}{(n-1)b^q} \sum_{l \neq j} (g(z_l) - g(z_j)) \bar{K}_{jl} \right] K_{ij} \\ &\triangleq \hat{\gamma}(z_i) + M_i + T_i \end{aligned}$$

where $M_i = \frac{1}{(n-1)a^p} \sum_{j \neq i} g(z_j) ((\tilde{f}_a(z_j) - f_a(z_j)) K_{ij})$ and $T_i = \frac{1}{(n-1)a^p} \sum_{j \neq i} \left[\frac{1}{(n-1)b^q} \sum_{l \neq j} (g(z_l) - g(z_j)) \bar{K}_{jl} \right] K_{ij}$. Hence $S_{\hat{g}-\hat{\gamma}} \equiv S_{M+T}$. By the Cauchy inequality, we only need to show that $S_M = o_p(n^{-1/2})$ and $S_T = o_p(n^{-1/2})$. We prove them below.

Proof of $S_M = O((nb^q)^{-1} + b^{2\mu})$:

$E|S_M| = \frac{1}{n} \sum_i E(M_i^2) = E(M_1^2)$. Using $M_1 = ((n-1)a^p)^{-1} \sum_{i \neq 1} g(z_i) (\tilde{f}_a(z_i) - f_a(z_i)) K_{i1} = [((n-1)a^p)^{-1} n^{-1} \sum_{i \neq 1} \sum_{j \neq i} g(z_i) K_{i1} (b^{-q} K_{ij} - f_a(z_i))]$, we have

$$\begin{aligned} E|S_M| &= E\left\{ [((n-1)a^p)^{-1} n^{-1} \sum_{i \neq 1} \sum_{j \neq i} g_i K_{1i} (b^{-q} K_{ij} - f_a(z_i))] \right. \\ &\quad \left. [((n-1)a^p)^{-1} n^{-1} \sum_{i' \neq 1} \sum_{j' \neq i'} g_{i'} K_{1i'} (b^{-q} \bar{K}_{i'j'} - f_a(z_{i'}))] \right\} \end{aligned} \quad (A.5)$$

We consider two different situations: (I) j and j' are both different from 1 and (II) at least one of the j and j' equals 1. The proof for case (I) is the same as in the proof of (A.3). For example, if i, j, i', j' are all different from each other, then using lemmas B.1 and B.3, (A.5) is $O(b^{2\mu})$, and if some of the i, j, i', j' take the same values, then (A.5) is of the order $O_p((nb^q)^{-1}) + o_p(b^{2\mu})$. The detailed proof is identical to the proof of (A.3) and is thus omitted here.

For case (II), The two subcases are symmetric, hence we only consider $j' = 1$, we still have different situations: (1) i, i', j take three different values, (2) i, i', j take two different values, (3) i, i', j all take the same value.

For (1), first assume $j \neq 1$, then (A.5) becomes (using lemmas B.1 and B.2)

$$\begin{aligned} & ((n-1)a^p)^{-2} n^{-2} \sum_{i \neq 1} \sum_{j \neq i, 1} \sum_{i' \neq i, i, j} E[g(z_i)g(z_{i'})K_{1i}(b^{-q}\bar{K}_{ij} - f_a(z_i))K_{1i'}(b^{-q}\bar{K}_{i'1} - f_a(z_{i'}))] \\ &= n^{-1} a^{-2p} E[g(z_2)g(z_3)(E_{1,z_2}K_{12})(E_{1,z_3}K_{13})(b^{-q}E_2(\bar{K}_{24}) - f_a(z_2))(b^{-q}E_{3,x_1}(\bar{K}_{31}) - f_a(z_3))] \\ &= n^{-1} E[g(z_2)g(z_3)G(z_2)G(z_3)O(b^{2\mu})] = O(n^{-1}b^{2\mu}) = o(b^{2\mu}) \end{aligned}$$

Next if $j = 1$, then (A.5) becomes

$$\begin{aligned} & ((n-1)a^p)^{-2} n^{-2} \sum_{i \neq 1} \sum_{i' \neq i, 1} E[g(z_i)g(z_{i'})K_{1i}(b^{-q}\bar{K}_{i1} - f_a(z_i))K_{1i'}(b^{-q}\bar{K}_{i'1} - f_a(z_{i'}))] \\ &= n^{-2} a^{-2p} E[g(z_2)g(z_3)(E_{1,z_2}K_{12})(E_{1,z_3}K_{13})(b^{-q}(\bar{K}_{21}) - f_a(z_2))(b^{-q}(\bar{K}_{31}) - f_a(z_3))] \\ &= n^{-2} O(1) E\{E_1[g(z_1)(b^{-q}(\bar{K}_{21}) - f_a(z_2))]E_1[g(z_3)(b^{-q}(\bar{K}_{31}) - f_a(z_3))]\} = O(n^{-2}). \end{aligned}$$

For cases (2) and (3), there are at most two summations and using the same arguments as above, it is easy to show that, in these cases, (A.5) will have a smaller order. Hence for case (II), (A.5) is of the order $O(n^{-1}b^{2\mu} + n^{-2})$. Summarizing the results for cases (I) and (II) above, we have shown that $E|S_M| = O((nb^q)^{-1} + b^{2\mu})$.

Proof for $S_T = O((nb^q)^{-1} + b^{2\mu})$:

$$\begin{aligned} & \text{Using } T_1 = [((n-1)a^p)^{-1}((n-1)b^q)^{-1} \sum_{i \neq 1} \sum_{j \neq i} K_{1i}(g(z_j) - g_a(z_i))\bar{K}_{ij}], \text{ we have } E|S_T| = \\ & E(T_1^2) = E\{[((n-1)a^p)^{-1} \sum_{i \neq 1} K_{1i}][((n-1)b^q)^{-1} \sum_{j \neq i} (g(z_j) - g_a(z_i))\bar{K}_{ij}][((n-1)a^p)^{-1} \sum_{i' \neq 1} K_{1i'}] \\ & [((n-1)b^q)^{-1} \sum_{j' \neq i'} (g(z_{j'}) - g_a(z_{i'}))\bar{K}_{i'j'}]\} \end{aligned} \quad (\text{A.6})$$

Comparing (A.6) and (A.5), the main difference is that $\tilde{f}_a(z_i) - f_a(z_i)$ in (A.5) is replaced

by $\tilde{g}(z_i) - g(z_i)\hat{f}_a(z_i)$. This amounts to replacing lemma B.2 by lemma B.3 in the proof. Thus, it is obvious that (A.6) has the same order as (A.5) and $S_T = O_p((nb^q)^{-1} + b^{2\mu})$.

Proposition 3 $S_{(\tilde{g}-\hat{g})f, uf} = o_p(n^{-1/2})$,

Proof: $S_{(\tilde{g}-\hat{g})f, uf} \equiv S_{F+H, uf}$. Let $A = F$ or H , we have $S_{A, uf} = S_{A, uf} + S_{A, u(f-f)}$. We will show that both $S_{A, uf}$ and $S_{A, u(f-f)}$ are $o_p(n^{-1/2})$. $E[(S_{A, uf})^2] = \frac{1}{n^2} \sum_i \sum_j E[(A_i u_i f_i)(A_j u_j f_j)] = \frac{1}{n^2} \sum_i E[A_i^2 u_i^2 f_i^2] = n^{-1} E[A_1^2 u_1^2 f_1^2] = n^{-1} O(E(A_1^2)) = n^{-1} O(E|S_A|) = n^{-1} O((nb^q)^{-1} + b^{2\mu})$ by proposition 1. $E|S_{u(f-f)}| = \frac{1}{n} \sum_i E[u_i^2 (\hat{f}_i - f_i)^2] = E[u_1^2 (\hat{f}_1 - f_1)^2] = O(E[(\hat{f}_1 - f_1)^2]) = O((na^p)^{-1} + a^{2\nu})$ by lemma B.4. Also $S_A = o_p(n^{-1/2})$ by proposition 1, hence $S_{A, u(f-f)} = o_p(n^{-1/2})$ by the Cauchy inequality.

Proposition 4 $S_{\hat{g}-\hat{r}, uf} = o_p(n^{-1/2})$,

Proof: $S_{\hat{g}-\hat{r}, uf} \equiv S_{M+T, uf} = S_{M+T, uf} + S_{M+T, u(f-f)}$. Let $B = M$ or T , then by a similar proof as proposition 3 (using proposition 2 and the Cauchy inequality), one can show that $S_{B, uf} = o_p(n^{-1/2})$.

Proposition 5 $S_{\tilde{r}f} = O_p(n^{-1/2})$, ($r = w$, or u).

We first consider the case of $r = w$. $S_{\tilde{w}f} = S_{\tilde{w}f} + S_{\tilde{w}(f-f)}$.

$$\begin{aligned} E|S_{\tilde{w}f}| &= \frac{1}{n} \sum_i E[\tilde{w}_i^2 f_i^2] = E[\tilde{w}_1^2 f_1^2] = E\{f_1^2 [((n-1)b^q)^{-1} \sum_{i \neq 1} w_i \bar{K}_{1i}] [((n-1)b^q)^{-1} \sum_{j \neq 1} w_j \bar{K}_{1j}]\} \\ &= E\{f_1^2 [((n-1)b^q)^{-2} \sum_{i \neq 1} w_i^2 \bar{K}_{1i}^2]\} = E\{f_1^2 [n^{-1} b^{-2q} E_1(w_2^2 \bar{K}_{12}^2)]\} = O((nb^q)^{-1}), \quad (A.7) \end{aligned}$$

by lemma B.1. We also used the fact that $E(w_i | Z, X_{-i}) = 0$, where $X_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.

Obviously (A.7) also proves $S_{\tilde{w}} = O_p((nb^q)^{-1})$ (by removing f_1^2). From $S_{\tilde{w}} = o_p(n^{-1/2})$ and $S_{\tilde{f}-f} = o_p(n^{-1/2})$ (see lemma B.4) and the the Cauchy inequality, we know that $S_{\tilde{w}(f-f)} = o_p(n^{-1/2})$.

The proof is identical to the case of $r = w$. Simply replacing w by u in the above proof.

Proposition 6 $S_{\hat{r}} = o_p(n^{-1/2})$, ($w = r$ or u).

Proof: $E|S_{\hat{w}}| = \frac{1}{n} \sum_i E[\hat{r}_i^2] = E[\hat{w}_1^2] = [(n-1)^4 a^{2p} b^{2q}]^{-1} \sum_{i \neq 1} \sum_{j \neq i} \sum_{i' \neq 1} \sum_{j' \neq i'} E[w_j w_{j'} \bar{K}_{ij} \bar{K}_{i'j'} K_{1i} K_{1i'}] = [(n-1)^4 a^{2p} b^{2q}]^{-1} \{ \sum_{i \neq 1} \sum_{j \neq i, i'} E[w_j^2 \bar{K}_{ij} \bar{K}_{i'j} K_{1i} K_{1i'}] + \sum_{i' \neq 1} (n-1)(n-1)(n-3) E[w_3^2 \bar{K}_{23} \bar{K}_{43} K_{12} K_{14}] + (n-1)(n-2) E[w_3^2 \bar{K}_{23}^2 K_{12}^2] \} = O(n^{-1} + (n^2 a^p b^q)^{-1}) = O(n^{-1})$. The same proof leads to $E|S_{\hat{u}}| = O(n^{-1})$.

Proposition 7 $S_{\bar{r}f,uf} = o_p(n^{-1/2})$, ($r = w$ or u).

Proof: $S_{\bar{r}f,uf} = S_{\bar{r}f,uf} = S_{\bar{r}f,u(f-f)}$. Then similar to the proof of proposition 3, one can show that $E|S_{\bar{r}f,uf}^2| = n^{-1} O(E[(\tilde{r}_1 \hat{f}_1)^2]) = o(n^{-1})$. Also $S_{u(f-f)} = o_p(n^{-1/2})$ (by the proof of proposition 3) and $S_{\bar{r}f} = o_p(n^{-1/2})$ by proposition 5. Hence $S_{\bar{r}f,u(f-f)} = o_p(n^{-1/2})$ by the Cauchy inequality.

Proposition 8 $S_{\hat{r},uf} = o_p(n^{-1/2})$, ($r = w$ or u).

Proof: $S_{\hat{r},uf} = S_{\hat{r},uf} + S_{\hat{r},u(f-f)}$. Then similar to the proof of proposition 7, using propositions 3, 6 and the Cauchy inequality, it is to see that $S_{\hat{r},uf} = o_p(n^{-1/2})$.

Proof of Theorem 2

Proof: Using similar arguments as in Theorem 1. It follows that, under H_a , $I_n \xrightarrow{p} E\{[m_a(z)f_a(z) - E(m_a(z)f_a(z)|x)][m_a(z) - E(m_a(z)|x)]f^2(x)\} \neq 0$, and $\hat{\sigma}_0^2 \xrightarrow{p} E\{(g(z)f_a(z) - E(g(z)f_a(z)|x))^2 f^2(x) + \epsilon^2 f^2(x)\}$, which is positive and finite. Hence $Prob[|J_n| > c] \rightarrow 1$ for any non-negative constant c .

Appendix B

In this appendix, we give some useful lemmas. Lemmas B.1 - B.3 are from Robinson (1988), and the proof of lemma B.4 can be found in Fan, Li and Stengos (1992). In the lemmas B.1-B.3 below, we assume that $x \in R^p$, $k \in \mathcal{K}_\nu$ (some $\nu \geq 1$), $f(v) \in \mathcal{G}_\nu^\infty$, $g(v) \in \mathcal{G}_\nu^\alpha$ for some $\alpha > 0$. E_1 denote the conditional expectation (conditional on x_1). K_{ij} denote $K(\frac{x_i - x_j}{a})$. Note that lemmas B.1-B.3 hold when changing (x, f, K, p, a, ν) to $(z, f_a, \tilde{K}, q, b, \mu)$.

Lemma B.1

$$(i) E_1(\hat{f}_1) = \frac{1}{(n-1)a^p} \sum_{i \neq 1} E_1(K_{1i}) = a^{-p} E_1(K_{12}) = O(1),$$

$$(ii) a^{-p} |E_1[(g(x_2)K_{12})]| \leq \Phi(x_1),$$

where $\Phi(x)$ has finite α th moment.

Lemma B.2

$$E_1(\hat{f}_1 - f_1) = E_1(((n-1)a^p)^{-1} \sum_{i \neq 1} K_{1i}) - f(x_1) = a^{-p} E_1(K_{12}) - f_1 \leq G(x_1)a^\nu,$$

where $G(\cdot)$ has finite moments of any order.

Lemma B.3

$$E_1(g_1 \hat{f}_1 - \hat{g}_1) = \frac{1}{(n-1)a^p} \sum_{i \neq 1} E_1(g_1 - g_i)K_{1i} = a^{-p} E_1[(g(x_1) - g(x_2))K_{12}] \leq D(x_1)(a^\nu),$$

where $D(\cdot)$ has finite α th moment.

Lemma B.4 Let m, ξ, f, f_a, u and v all defined as in the paper. Let $g = m$ or ξ and $\epsilon = u$ or v , then

$$(i) S_{g\hat{f}-\hat{g},\epsilon\hat{f}} = O_p((na^p)^{-1}), S_{g\hat{f}-\hat{g},\epsilon} = O_p((na^p)^{-1}),$$

$$(ii) S_{g\hat{f}-\hat{g}} = O_p(((na^p)^{-1} + a^{2\mu})), S_{\hat{f}-f} = O_p(((na^p)^{-1} + a^{2\mu})),$$

$$(iii) S_{v,\hat{u}} = O_p(((na^p)^{-1})), S_{\hat{v},uf} = O_p(((na^p)^{-1})), S_{\hat{v},\hat{u}} = O_p(((na^p)^{-1})).$$

References

- Cox, D.R. (1961) "Tests of separate families of hypotheses" *Proceedings of the Fourth Berkeley Symposium of Mathematical Statistics and Probability*, 1, 105-123.
- Cox, D.R. (1962) "Further results on tests of separate families of hypotheses" *Journal of the Royal Statistical Society, Series B*, 24, 406-424.
- Davidson, R. and J.G. MacKinnon (1981) "Several model specification tests in the presence of alternative hypotheses" *Econometrica* 49, 781-793.
- Davidson, R. and J.G. MacKinnon (1982) "Some non-nested hypothesis tests and the relations among them" *Review of Economic Studies* 49, 551-565.
- de Jong, R.M. and H.J. Bierens (1994) "On the limit behavior of a Chi-Square type test if the number of conditional moments tested approaches infinity" *Econometric Theory* 10, 70-90.
- Delgado, M.A. and T. Stengos (1994) "Semiparametric specification testing of non-nested econometric models" *Review of Economic Studies* 61, 291-303.
- Eubank, R.L. and C.H. Spiegelman (1990) "Testing the goodness of fit of a linear model via non-parametric regression techniques" *Journal of the American Statistical Association* 85, 387-392.
- Fan, Y., Q. Li and T. Stengos, (1992) "Root-N-consistent semi-parametric regression with conditionally heteroskedastic disturbances" *Journal of Quantitative Economics*, forthcoming.
- Fan, Y., Q. Li, (1994) "Consistent non-parametric model specification tests: Omitted variables, parametric and semi-parametric functional forms." Working paper, Department of Economics, University of Guelph.
- Fisher, G.R. and M. McAleer, (1981) "Alternative procedures and associated tests of signifi-

cance for non-nested hypotheses" *Journal of Econometrics* 16, 103-119.

Godfrey, L.G. and M.H. Pesaran (1983) "Tests of non-nested regression models: small sample adjustments and Monte Carlo evidence" *Journal of Econometrics* 21, 133-154.

Hall, P. and J.D. Hart (1990) "Bootstrap test for difference between means in non-parametric regression" *Journal of the American Statistical Association* 85, 1039-1049

Härdle, W. and E. Mammen (1993) "Comparing non-parametric versus parametric regression fits" *The Annals of Statistics*, 21, 1926-1947.

Hong and White (1991) "Consistent specification testing via non-parametric series regressions" Working paper, Department of Economics, University of California, San Diego.

Horowitz J.L. and W. Härdle, (1994) "Testing a parametric model against a semi-parametric alternative" mimeo.

Pesaran, M.H. (1974) "On the general problem of model selection" *Review of Economic Studies* 41, 153-171.

Powell, J.L., J.H. Stock, and T.M. Stocker, (1989), "Semiparametric estimation of the index coefficients" *Econometrica* 57, 1043-1430.

Robinson, P., (1988) "Root-N consistent semi-parametric regression" *Econometrica* 156, 931-954.

Whang, and D. Andrews (1993) "Tests of Specification for parametric and semi-parametric models" *Journal of Econometrics*, 277-318.

Wooldridge, J. (1992) "A test for functional form against non-parametric alternatives " *Econometric Theory* 8, 452-475.